

SÓLO PARA PARTICIPANTES

DOCUMENTO DE REFERENCIA
DDR/15
24 de abril de 2001

ESPAÑOL
ORIGINAL: PORTUGUÉS

CEPAL

Comisión Económica para América Latina y el Caribe

Primera reunión de la Conferencia Estadística de las Américas
de la Comisión Económica para América Latina y el Caribe

Santiago de Chile, 9 al 11 de mayo de 2001

DESAFÍOS Y TENDENCIAS DE LA MODERNIZACIÓN DE LOS SISTEMAS ESTADÍSTICOS NACIONALES

El tratamiento de la confidencialidad de las informaciones estadísticas en el
Instituto Brasileño de Geografía y Estadística

Este documento fue preparado por Zelia Magalhães Bianchini, Directora adjunta de investigación y coordinadora del Grupo de Confidencialidad del Instituto Brasileño de Geografía y Estadística (IBGE). Las opiniones expresadas en este documento, que no ha sido sometido a revisión editorial, son de la exclusiva responsabilidad de la autora y pueden no coincidir con las de la Organización.

01-2-149

1. Introducción

El **objetivo del Instituto Brasileño de Geografía y Estadística (IBGE)** fue siempre y sigue siendo proporcionar información relevante y adecuada a la sociedad. Satisfacer ese objetivo implica proveer datos con tanta riqueza y detalle como los recursos lo permitan. A su vez, el hecho de aumentar la riqueza y los detalles implica el riesgo de que uno o más de los datos producidos y difundidos conduzcan a la revelación de datos sobre individuos, entidades y empresas identificados o identificables. En ese contexto, están en juego los **derechos** de la sociedad a la información y los derechos de los individuos, entidades y empresas a la privacidad de las informaciones individuales.

Los **avances tecnológicos** ofrecen oportunidades para almacenar, procesar, acceder y analizar grandes conjuntos de datos en forma eficiente, lo que refuerza la exigencia de los usuarios por informaciones cada vez más detalladas. Por otro lado, facilitan el acceso a datos identificables y confidenciales, aumentando el riesgo de su revelación individual. La **legislación** que confiere al IBGE el mandato para la producción de estadísticas es explícita en el sentido de establecer la obligatoriedad de la prestación de informaciones al IBGE, pero a su vez impone como contrapartida que las informaciones entregadas tendrán carácter secreto, serán usadas exclusivamente para **finestadísticos** y no podrán ser objeto de certificación ni servirán de prueba en un proceso administrativo, fiscal o judicial ... (véase la **Ley 5534**, de **14/11/68** y el **Decreto-ley 161**, de **13/02/67**).

La tarea de encontrar el equilibrio entre lo que se solicita y lo que debe ser accesible o permitido no resulta fácil.

Ante la necesidad emergente de crear **normas y procedimientos** destinados a la mantención de la confidencialidad de la información recopilada, producida, difundida y almacenada por el IBGE, se creó en 1999 el **Grupo de la Confidencialidad**, cuyas metas se concentraron en la recomendación de líneas de acción, tendientes a **reducir los riesgos de revelación y observar el cumplimiento de la promesa de asegurar la privacidad de las informaciones confidenciales**, a través del anonimato y la minimización de los riesgos de revelación.

El Grupo de la Confidencialidad trató de operar con transparencia y visibilidad para todos, pero sobre todo tratando de conquistar la responsabilidad y el compromiso de todos los que participan en el proceso de producción de información o que, casualmente, fueran autorizados a tener acceso a las informaciones individuales. Tales líneas de acción, amparadas en la legislación, la práctica de la abrumadora mayoría de los países, las recomendaciones del Instituto Internacional de Estadística, y la propia experiencia del IBGE de más de cincuenta años en la producción de estadísticas oficiales de calidad son **esenciales/fundamentales para la continuación de la existencia del IBGE como institución pública digna de la fe pública**, capaz de prestar servicios con imparcialidad, integridad y calidad.

La integridad de las estadísticas oficiales y la confianza pública en esas estadísticas son esenciales para el buen gobierno de una nación. Para que exista esa confianza pública en las

estadísticas oficiales, es esencial asegurar la **integridad de las instituciones** encargadas de su producción, como es el caso del IBGE, que depende de manera crucial de la confianza pública para obtener las informaciones que necesita para entregar al gobierno y la sociedad las estadísticas (oficiales) necesarias para la información, el debate y la toma de decisiones.

Una de las condiciones esenciales para mantener la integridad y la fe pública de una institución productora de estadísticas como el IBGE (y como todos sus homólogos en el mundo) es la **preservación intransigente de la confidencialidad de las informaciones** individuales o identificadas que utiliza como parte de su proceso de producción de estadísticas. Por definición, las estadísticas son construcciones basadas en agregar informaciones individuales, quitándoles la individualidad y la identidad, y construyendo resúmenes de las características relevantes del conjunto de individuos, empresas, entidades, productos, etc.

La cuestión de la mantención de la confidencialidad de las informaciones debe abordarse según sus diversos aspectos relevantes, que comprenden el aspecto legal, el político, el ético y uno de carácter técnico-operacional. Resulta peligroso omitir cualquiera de estos aspectos, pues ello puede conducir a decisiones cuyo efecto será desastroso para la credibilidad de la Institución.

El presente texto ofrece una descripción de los principales aspectos analizados en el ámbito de las actividades realizadas por el Grupo de la Confidencialidad, que incluyen los desafíos y las perspectivas para el futuro.¹ Se presentan las líneas de trabajo, la operacionalización (entidades y actores que actúan en la producción de informaciones), y la situación actual de cada línea de trabajo (métodos, compromisos, responsabilidades, caminos de acceso a la información, ley estadística y ética). Se identifican muchas acciones: investigar, estudiar, concientizar, diagnosticar, reexaminar la legislación, organizar directrices, documentar, automatizar procedimientos, orientar, recomendar, supervisar, preparar normas, etcétera.

Habida cuenta de lo anterior, e identificando la oportunidad para analizar el problema a fondo con miras a la formulación de normas y procedimientos destinados a mantener la confidencialidad de la información recopilada, almacenada, producida y difundida por el IBGE, se procuran abordar en las secciones siguientes los diferentes aspectos de la cuestión.

En la **sección 2** se encuentran las líneas de trabajo que se han adoptado y un flujo con el **modelo de operación** donde se identifican las entidades y los agentes involucrados que actúan en cada punto del flujo de la producción de información y lo que debe observarse en cada uno de esos puntos (en lo que respecta a la confidencialidad). Este flujo debe orientar la preparación de las recomendaciones con miras a una política sobre la confidencialidad de la información para el IBGE.

La **sección 3** presenta, en forma bastante resumida, los aspectos principales que se tratan con miras a la **implementación de ese modelo operacional** a través de los conceptos básicos, los métodos y técnicas de protección de la información, la legislación competente y los procedimientos de conducta ético-moral. Presenta, también, una propuesta de control del acceso al acervo institucional de datos.

¹ Véase Bianchini y otros, 1999.

En la **sección 4** se indican **los desafíos y las perspectivas para el futuro** y, por último, en la **sección 5**, se presenta una vasta bibliografía disponible respecto del tratamiento de la confidencialidad de la información.

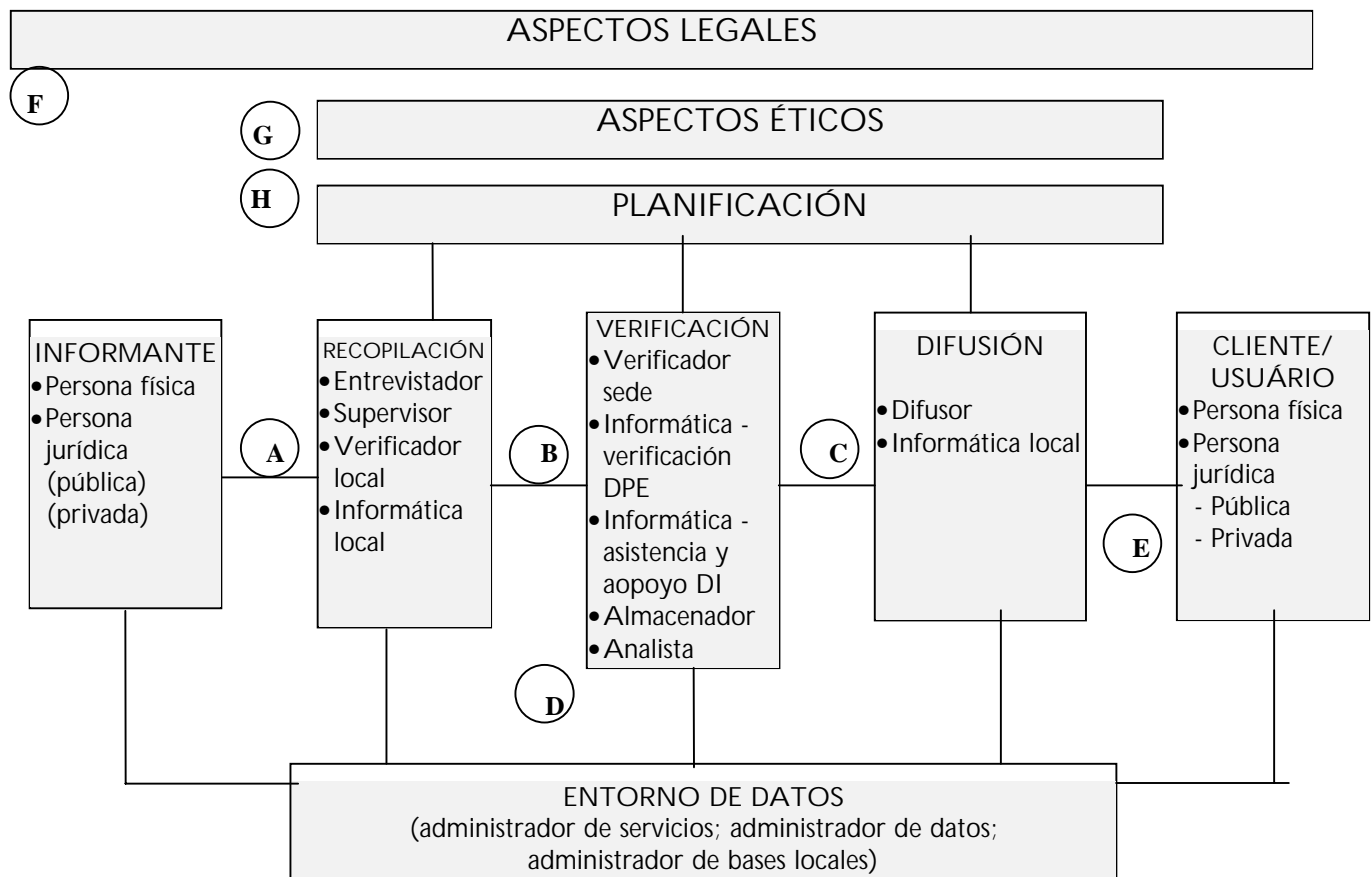
2. Líneas de trabajo y modelo de operación

Las **principales líneas de trabajo** detectadas en el ámbito de las actividades realizadas fueron las siguientes:

- **Investigación bibliográfica de las prácticas internacionales** de los principales organismos productores de información para tratar la cuestión de la confidencialidad, como complemento para la elaboración de recomendaciones para tratar la cuestión en el IBGE.
- **Elaboración de una ley estadística** para el país, mediante la revisión, actualización, ampliación y **consolidación** de la legislación vigente, para otorgarle actualidad democrática y actualizar la ley a la época tecnológica en que vivimos.
- **Definición conceptual y diseño de métodos y técnicas** apropiados para minimizar los riesgos de revelar la **información** confidencial en los datos preparados para ser difundidos, sean estos datos agregados/tabulados, microdatos, datos de directorios, **datos** georreferenciados, etcétera.
- **Definición de los aspectos técnico-operacionales** que se incluirán en manuales, de forma de homogeneizar el **tratamiento** de la cuestión de la confidencialidad en **las** diversas áreas del IBGE, no sólo respecto de los medios y técnicas empleados sino también de los resultados que se obtienen.
- **Implementación rutinaria del acceso al acervo de datos** no preparados para ser difundidos. Todo acceso **debe** ser autorizado/cancelado con arreglo a criterios claramente establecidos, aplicables a todos los funcionarios (activos e inactivos) o **personas** vinculadas con entidades externas con las cuales el IBGE mantiene convenios y/o contratos de prestación de servicios.
- **Definición de los principios éticos** que deben inspirar la postura/conducta de los funcionarios. Tales principios deben considerar tres vertientes, a saber:
 1. El cumplimiento de la ley estadística;
 2. El cumplimiento de los procedimientos técnico-operacionales existentes, y
 3. El tratamiento de los casos de negligencia.
- **Difusión del tema** a través de seminarios de concientización, creación de la dirección electrónica para el envío de comentarios y sugerencias de los **interesados** en contribuir a los trabajos del grupo, y presentación de las actas de reuniones y las referencias bibliográficas disponibles en medio digital en la Intranet.

A fin de sintetizar y orientar la concreción de estas líneas de trabajo y facilitar la planificación de los trabajos, la comunicación entre sus miembros y de éstos con los demás funcionarios de la institución que participan en la recopilación, producción, almacenamiento y difusión de la información se estructuró el **modelo de operación** que se reseña. Para mayores detalles, véase Anzanello (1999).

MODELO DE OPERACIÓN Y CONFIDENCIALIDAD DE LA INFORMACIÓN



Este modelo de operación procura abarcar en una visión integrada los diversos aspectos de la cuestión de la confidencialidad (legales, éticos y operacionales), **tratando de identificar la relación entre las diversas entidades y agentes involucrados** que actúan en cada punto del proceso de producción de la información y lo que debe observarse (respecto de la confidencialidad) en cada uno de ellos.

La cuestión del almacenaje de datos está ligada a las diversas etapas del flujo de la información, desde la recopilación, pasando por la verificación y difusión, hasta llegar al cliente/usuario. Hay varias preguntas que deben responderse:

- ¿Cómo almacenar los datos?
- ¿A quién dar acceso?

- ¿En qué forma dar acceso?
- ¿Cómo controlar el acceso y uso?

La cuestión de la seguridad abarca también todo el recorrido de la información. De nada sirve, por ejemplo, tener la máxima seguridad respecto al entorno de datos en materia de verificación y difusión, si no existe la misma seguridad en la etapa de la recopilación.

Las letras contenidas en los círculos del modelo de operación indican los **puntos del flujo de la información que deben considerarse en la cuestión de la confidencialidad**. En cada punto de observación se destacan:

- La relación entre entidades del modelo;
- Los agentes involucrados que actúan en cada punto del recorrido de la información;
- Lo que se debe observar (objetos relacionados con las cuestiones de la confidencialidad de la información), y
- Orientaciones (para cada objeto que se considere).

Para cada uno de estos puntos debemos tratar de responder algunas preguntas básicas relacionadas con la confidencialidad de la información:

- 1) ¿Quién tiene acceso habitual a la información secreta/confidencial y con quién se relaciona durante la realización de su trabajo?
- 2) ¿En esas relaciones, como garantizar que no se viole la confidencialidad?
- 3) ¿Quién autoriza a las personas a ocuparse de la información confidencial contenida en el papel? ¿Existen instrucciones sobre cómo proceder? ¿El acceso es controlado? ¿Hay sanciones si se viola la confidencialidad?
- 4) ¿Quién autoriza a las personas a ocuparse de la información confidencial procesada y almacenada en forma digital? ¿Existen instrucciones sobre cómo proceder? ¿El acceso es controlado? ¿Hay sanciones si se viola la confidencialidad?
- 5) ¿Cómo ocuparse de los datos contenidos en el material impreso (manipulación, almacenamiento, recuperación, tratamiento, transporte, descarte)? ¿Existen instrucciones sobre cómo proceder?
- 6) ¿Cómo ocuparse de los datos almacenados en forma digital (manipulación, almacenamiento, copia, recuperación, tratamiento, transmisión, descarte)? ¿Existen instrucciones sobre cómo proceder?

Como complemento, definir las líneas de acción que deben implementarse para resolver casos concretos en cada uno de los puntos de observación.

3. Implementación del modelo de operación

En esta sección se reseñan los principales aspectos considerados.

Respecto a la concepción y métodos para limitar la revelación (violación de la confidencialidad) en datos preparados para ser difundidos, se consideró como referencia básica el texto titulado *Statistical Policy Working Paper 22-Report on Statistical Disclosure Limitation Methodology*, Estados Unidos, 1994, preparado por el *Subcommittee on Disclosure Limitation Methodology* y coordinado por Nancy J. Kirkendall.

3.1. Conceptos básicos

Se destacan la concepción sobre "**¿Qué es revelación?**" y los tipos de revelación: revelación de la identidad, revelación de atributos y revelación deductiva.

La **revelación de la identidad** ocurre cuando un tercero puede identificar a un declarante a partir de los datos divulgados. En el ámbito de los microdatos, la identificación se considera generalmente como revelación. Por tal motivo, los métodos para limitar la revelación aplicables a los micro datos limitan o modifican la información que podría utilizarse para identificar declarantes específicos o unidades de referencia.

La **revelación de atributos** ocurre cuando se revela una información confidencial sobre el individuo y puede atribuirse al mismo. La revelación de atributos presupone que hubo una revelación de la identidad. Las técnicas para limitar la revelación aplicables a las tablas aseguran que los datos del declarante son publicados sólo como parte de un agregado, con un número suficiente de otros declarantes con el fin de prevenir la revelación de atributos.

La **revelación deductiva** ocurre cuando la información puede deducirse con alta confiabilidad a partir de las propiedades estadísticas de los datos divulgados. Por ejemplo, los datos pueden mostrar una alta correlación entre el ingreso y el gasto del hogar. En el caso de que éste último sea una información pública, un tercero puede usar esta información para deducir el ingreso. Las deducciones están destinadas a predecir el comportamiento agregado y no los atributos individuales.

A partir de los clasificadores y atributos se debe desarrollar la técnica para impedir la revelación de la identidad en las diversas formas de acceso a los datos: tabulaciones básicas, tabulaciones especiales, microdatos de encuestas domiciliarias, acceso asistido (micro datos del área económica), Internet, etcétera.

3.2. Métodos y técnicas

3.2.1 Protección de los datos tabulados

La selección de una regla de la confidencialidad estadística para los datos presentados en las tablas (datos tabulados) depende de si los datos son frecuencias u otras cantidades pertinentes que deben medir algo más que el recuento de unidades en la celda de la tabla.

En el caso de las **tablas de frecuencias**, los métodos principales para proteger la confidencialidad son los siguientes:

- **Muestreo y estimación de las frecuencias en las celdas**
- **Reglas especiales** - imponen restricciones al nivel de detalle que puede darse a una tabla, por ejemplo, prohíben las tabulaciones en las cuales una determinada celda es igual al total marginal.
- **La regla del umbral** - una celda de la tabla de frecuencias se define como de riesgo si el número de declarantes es menor que un número específico determinado. Para reducir el riesgo de revelación, las tablas pueden reestructurarse y las categorías combinarse o pueden utilizarse métodos como la supresión de celdas, el redondeo aleatorio, el redondeo controlado o la edición de confidencialidad:
 - **Supresión de celdas** - en un renglón o columna con una celda de riesgo suprimida, debe suprimirse por lo menos otra celda (supresión complementaria);
 - **Redondeo aleatorio** - en vez de utilizar el redondeo patrón, se toma una decisión aleatoria indicando si el valor se redondeará hacia arriba o hacia abajo;
 - **Redondeo controlado** - es una forma de redondeo aleatorio, pero con la limitante de que la suma de las entradas publicadas en cada renglón y columna tiene que ser igual al total marginal publicado correspondiente;
 - **Edición de confidencialidad** - técnica de limitación de la revelación estadística que puede aplicarse a los archivos de microdatos antes de que éstos sean utilizados en la preparación de las tablas.

Las tablas que contienen datos cuantitativos tienen un problema de revelación único. Las distribuciones de estos valores son usualmente asimétricas, con pocas unidades que tienen valores muy grandes. En este caso, la limitación de la revelación se centra en evitar que los datos publicados se utilicen para estimar con precisión los valores de las unidades más grandes. Al proteger los valores más grandes, estaremos protegiendo de hecho todos los valores.

Se han elaborado reglas para las tablas de datos cuantitativos, denominadas reglas de supresión primaria o medidas de riesgo lineal, que determinan si una celda dada de la tabla podría revelar informaciones individuales del declarante. Tales celdas se denominan celdas de riesgo y no pueden publicarse. Una vez identificadas las celdas de riesgo hay dos opciones: reestructurar la tabla y juntar las celdas hasta que se acaben las celdas de riesgo o suprimir celdas.

Varios organismos recurren a una fórmula administrativa para evitar la supresión de celdas. Obtienen un permiso por escrito de los declarantes que contribuyen a las celdas de riesgo, que autoriza la publicación de las mismas. Ese permiso se conoce como la renuncia a la promesa de proteger el secreto de la información individual.

La **práctica corriente en el IBGE** para tratar la no identificación de datos tabulados para ser divulgados se da, tradicionalmente, sólo en las encuestas del área económica. El método utilizado es la regla del umbral, donde se exigen por lo menos tres declarantes por celda, donde cabe destacar que:

- sólo las encuestas económicas en materia de industria, comercio, servicios y existencias (de la agricultura) adoptan esta práctica en el sentido de no entregar información para celdas con menos de tres informantes, con el uso de técnicas específicas para este fin;
- tradicionalmente los procedimientos en esta línea de no identificación se aplicaban en la divulgación de los resultados de los censos económicos; se usaba el procedimiento de eliminar información en las celdas con menos de tres informantes (mediante una X); las encuestas anuales no empleaban el mismo procedimiento porque dado su diseño no muestral sólo divulgaban información en niveles espaciales y de clasificación de actividades más agregados, sin que se planteara por lo tanto el problema de las celdas con menos de tres informantes;
- el método general que se emplea actualmente tanto en la encuesta industrial anual (PIA) como en la encuesta anual del comercio (PAC) es el del umbral con por lo menos tres informantes en la celda y la opción preferencial por el procedimiento de agregación de niveles de clasificación de actividades (o espacial) dentro del objetivo de menor supresión de información; en la aplicación de estas reglas se opera con dos variantes: en las encuestas de la industria la agregación dentro de las agrupaciones de la clasificación nacional de actividades económicas (CNAE) sigue el criterio de selección de clase/grupo/división al que se agrega la celda con menos de tres informantes en función del menor valor de transformación industrial (VTI), dado el hecho de que la clasificación de la industria tiene más entradas y, por lo tanto, sus subdivisiones son más homogéneas; en la encuesta del comercio, la agregación dentro de las agrupaciones de la CNAE obedece al criterio de homogeneidad entre las actividades, pues la clasificación del comercio tiene menos entradas de dos y tres dígitos, lo que la hace más heterogénea en su interior;
- respecto a la cuestión del tratamiento de la no identificación en las encuestas del área económica, se trabaja con dos perspectivas de tiempo, a corto plazo, dentro de la línea ya expuesta realizándose el trabajo manualmente, y a mediano plazo invirtiendo en el sentido de utilizar programas computacionales para facilitar la labor de desidentificación.

El problema con los procedimientos actuales de tratamiento manual es el enorme volumen de trabajo y tiempo, lo que puede llegar a hacer inviable el despacho más amplio y rápido de las tabulaciones especiales de las encuestas económicas, recordando que en el caso de estas encuestas es la única forma de servir a los usuarios, toda vez que no se cuenta con microdatos para estas encuestas tal como ocurre con las encuestas domiciliarias.

3.2.2. Protección de los microdatos difundidos

Por un lado, la realidad impone una demanda cuya naturaleza implica proveer microdatos con una geografía detallada y por otro, el mandato de la confidencialidad de la información individualizada de la que es depositario el IBGE conduce a la restricción de datos, sea por agregación, eliminación o modificación o por restricción del acceso a través del control, la limitación del acceso o restricción de las vinculaciones. El desafío consiste en tratar de prestar los servicios que involucren microdatos sin incurrir en el riesgo de que uno o más de los datos producidos conduzcan a la revelación de información sobre individuos, entidades, empresas identificadas o identificables.

Los tipos de datos presentes en **los archivos de microdatos pueden clasificarse como identificadores, atributos de dominio público (clasificadores) y confidenciales**, diferenciándose de acuerdo con el tipo de unidad encuestada: domicilio y/o persona, unidad económica (empresa, establecimiento, etcétera), producto y precio.

Algunos microdatos incluyen **identificadores explícitos, tales como nombre y dirección**. La remoción de tales identificadores es obviamente el primer paso preparatorio para liberar un archivo cuya confidencialidad de la información individualizada debe ser protegida.

Además, hay dos factores que contribuyen a aumentar los riesgos de revelación. El primero, es la existencia de **información con alta visibilidad**. Algunos registros pueden representar declarantes con una característica única no local, tales como determinadas ocupaciones en ciertos municipios (juez, padre, médico, etc), o con rendimientos muy altos. El segundo, es la posibilidad de que **el archivo de microdatos se superponga con otros archivos** con informaciones más detalladas.

Es muy difícil proteger de la revelación a un archivo de microdatos debido a la posibilidad de **asociación con fuentes de datos externas** (Bethlehem, Keller y Pannekoek, 1990). Además, no hay medidas aceptables del riesgo de revelación para un archivo de microdatos, y por lo tanto no hay un patrón que pueda aplicarse para garantizar que la protección es adecuada.

Los métodos para reducir el potencial de revelación limitan o modifican la información. En general, los archivos de microdatos de uso público:

- incluyen sólo los datos de una muestra de la población;
- no incluyen identificadores obvios;
- limitan el detalle geográfico; y
- limitan el número de variables en el archivo.

Los métodos adicionales que se utilizan para ocultar las variables más evidentes comprenden:

- codificación excesiva o insuficiente;
- recodificación por intervalos o redondeos;
- adición o multiplicación por números aleatorios (ruidos);
- permutación de puestos (denominado también "switching");
- selección aleatoria de registros, atribuyéndoles un valor en blanco a algunas variables seleccionadas e imputándolas nuevamente;
- agregación de pequeños grupos de declarantes, sustituyendo el valor informado por cada individuo por la media del grupo (denominado también "blurring").

La práctica corriente del IBGE para difundir microdatos de las encuestas domiciliarias es a través de archivos de uso público, sin identificadores, con geografía poco detallada, pero sin tratamiento para ocultar las variables más evidentes.

En las encuestas económicas, la práctica de los principales órganos internacionales de estadística es no liberar el archivo de microdatos. Los riesgos de revelación en microdatos de encuestas del área económica son mucho más elevados que aquellos para los microdatos de las encuestas domiciliarias. Hay un mayor número de variables de identificación.

En general, las informaciones del área económica son datos cuantitativos. Estos datos son habitualmente asimétricos y permiten que los declarantes sean identificados fácilmente a partir de otras informaciones abiertas al público, además, pueden contener informaciones estratégicas. Por estas razones los microdatos del área económica no se difunden para uso público.

Ante la gran demanda de acceso a los microdatos del censo agropecuario, el IBGE viene tratando caso por caso las solicitudes de atención de varios órganos, autorizando el acceso in situ, es decir, el acceso a los microdatos en el Departamento de Agricultura mediante la presentación del proyecto, y la suscripción de un acuerdo para cada caso específico, en que se establecen las condiciones de uso y responsabilidad y se garantiza expresamente la confidencialidad. Los aspectos principales contemplados en el acuerdo para acceder a esos microdatos son los siguientes:

- identificación del proyecto en que se utilizarán;
- plazo de realización del proyecto;
- compromiso de utilizar los microdatos sólo en las dependencias del IBGE;
- observancia de la confidencialidad conforme a la legislación pertinente;
- compromiso de que las informaciones individuales no serán traspasadas, comercializadas o transferidas;
- compromiso de presentar los resultados para su liberación;
- en caso de que el producto final sea presentado en forma de tabla, adopción de procedimientos de no identificación previos a la liberación.

Como fruto de este esfuerzo de tratar de satisfacer las nuevas necesidades de los usuarios en la mejor forma posible, respetando el mandato de la confidencialidad, se están definiendo algunas orientaciones. Una de ellas se refiere a la **creación de espacios y el desarrollo de infraestructura y mecanismos de control y seguimiento**, con el nivel de seguridad adecuado, para otorgar acceso a los microdatos identificados o susceptibles de identificación trivial. A propósito, Zayatz, Masell y Skel (1999) mencionan la creación de entornos especiales, por parte de la Oficina del Censo de los Estados Unidos, para que los investigadores accedan a datos no liberados para su difusión.

3.2.3. Protección de microdatos con identificadores (directorios)

El tema de la confidencialidad en materia de directorios es bastante amplio, en función de los diversos directorios existentes: directorio de empresas, de establecimientos agrícolas, de domicilios (mencionados en las hojas de recopilación de los censos demográficos), de productos y precios. En los directorios el problema de revelación es el de la identificación, a diferencia de los microdatos, cuando están asociados los atributos.

Directorio de empresas

En primer lugar, se tratará específicamente el directorio orientado al área económica, es decir, el **directorio de empresas-CEMPRE**. La División de Directorio y Clasificación de la Dirección de Encuestas es la dependencia encargada de la preparación y mantención del CEMPRE, actualizándolo con los resultados provenientes de las encuestas continuas y la información proveniente de registros administrativos producidos externamente.

La situación actual sobre el tema de la confidencialidad en el **directorio de empresas** es la siguiente:

- La preocupación por la confidencialidad es pertinente cuando se trata de suministrar informaciones del directorio, con la identificación de nombres, direcciones y características de las empresas.
- El IBGE suministra selectivamente el directorio, entregando sólo el nombre, la dirección, la actividad y los grupos etarios del personal ocupado, a los usuarios que presentan justificaciones para su empleo. Se evita proporcionarlo para fines comerciales, del tipo catálogo postal.
- Hay consenso en cuanto a la provisión de directorios con las características de las empresas (por ejemplo: personal ocupado e ingreso) a los organismos regionales de estadística (OREs), con la finalidad de seleccionar muestras, previo acuerdo de resguardar la confidencialidad.
- El directorio no tiene finalidad comercial, su información no es estadística, pues ésta se caracteriza por la agregación de datos.
- La función principal del directorio es servir de instrumento de apoyo para la planificación de las encuestas económicas. El producto del IBGE está en los resultados de sus encuestas. Otra función del directorio está en su utilización como instrumento de coordinación estadística. El problema no es ceder el directorio a las empresas públicas o privadas, sino limitar su uso al diseño de encuestas.
- Otros aspectos se refieren a la calidad de las informaciones del directorio, la que puede cuestionar la credibilidad del organismo de estadística. El Statistics Canada al no entregar la información del directorio tiene presente la calidad de esa información.
- Los organismos de estadística de otros países, como el de Australia, que antes suministraban datos individualizados del directorio, informan en su página web que no proporcionan datos individualizados del directorio, sino sólo tabulaciones. Por otro lado, se tiene conocimiento de que el Instituto Nacional de Estadística y Estudios Económicos (INSEE) de Francia y el Instituto Nacional de Estadística (ISTAT) de Italia no tienen restricciones para proporcionar directorios.

Directorio de domicilios, establecimientos agropecuarios y de segmentos de espacios públicos

A diferencia del CEMPRE, los **directorios de domicilios y establecimientos agropecuarios** no están aún estructurados en formato digital. La fuente principal de informaciones para la estructuración de estos directorios son las hojas de recopilación llenadas con ocasión del censo demográfico² y del censo agropecuario, respectivamente, insertas en las libretas censales. Además, para cada encuesta domiciliaria se realiza la operación de listado, donde para todos los sectores seleccionados en cada muestra se prepara un directorio para efectuar la selección de domicilios.

El IBGE viene atendiendo la solicitud de provisión de copias de las hojas de recopilación o del listado, proporcionando la dirección y quitando el nombre y todos los demás atributos.

Se encuentra en vías de estructuración un **directorio de segmentos de espacios públicos** por sector que utiliza las informaciones de las hojas de recopilación del recuento de población de 1996: los nombres de los espacios públicos contenidos en el sector, el número de unidades visitadas y de domicilios particulares. Este directorio fue de gran utilidad para hacer la revisión de la base geográfica para el censo 2000. Este directorio se actualizará con las informaciones de las hojas de recopilación del censo 2000.

Directorio de productos y precios

Tradicionalmente, el IBGE sólo proporciona **precios medios** para productos genéricos, sin identificación de marca. A pesar de la fuerte demanda para que se entreguen los precios medios de determinados productos, por marca, el IBGE viene respondiendo negativamente basado en la garantía del secreto, la protección del informante y la preservación de la imagen de la institución.

La raíz de la dificultad que tiene el IBGE para satisfacer estos pedidos está precisamente en el concepto de lo que es una estadística, y de como entra en dicho proceso la información recopilada por el IBGE, pues no se trata de medias anónimas sino de medias referentes a productos y marcas bien especificados y por tanto identificados. Esa asociación entre el dato y una identificación es lo que causa el problema, lo que ha llevado al IBGE a no satisfacer pedidos de esta naturaleza.

3.2.4. Protección de datos georeferenciados.

El grupo examinó el problema de la revelación estadística en la construcción de **áreas pequeñas**, basándose en los artículos de Rees y Duke-Williams (1997 y 1998), preparados en el ámbito del proyecto de control de la revelación estadística con recursos del programa ESPIRIT de la Unión Europea.

Para evitar la revelación, una de las protecciones que se aplican es la definición de un límite mínimo de personas o domicilios para generar las estadísticas de las áreas. Si las estadísticas se generan para dos bases geográficas, entonces es posible, a través de operaciones del Sistema de Informaciones Georeferenciadas (SIG) sustraer un conjunto de tablas de otro para obtener

² Se ha planificado la transcripción a formato digital de las informaciones del censo demográfico 2000.

estadísticas de áreas bajo el límite de confidencialidad. El peligro de la diferenciación ocurre cuando es posible anidar áreas pequeñas de una base geográfica en áreas de otra base geográfica.

Estos experimentos sugieren que los riesgos prácticos son mucho menores de los que preconiza la teoría y que resulta seguro publicar estadísticas para bases geográficas secundarias en la medida en que éstas sean suficientemente más grandes que las áreas primarias.

Las recomendaciones de Rees y Duke-Williams (1997) para evitar el problema de la revelación, son las siguientes:

- Publique estadísticas de áreas pequeñas para los censos decenales. Existe una demanda importante al respecto y representa una estrategia de costo más efectiva que el servicio de tabulaciones por encargo.
- Seleccione la estrategia de producción de tablas seguras. Utilice los niveles mínimos de límite de confidencialidad para las tablas.
- Convierta la base geográfica primaria seleccionada en el uso más general posible.
- El riesgo de publicar en bases geográficas alternativas cuyas zonas sean mayores que las unidades censales primarias es ínfimo.
- Tal vez no sea seguro producir un conjunto secundario de estadísticas de áreas pequeñas de otra base geográfica minúscula, que esté relativamente próxima a la base geográfica primaria seleccionada. Esto rige para las publicaciones de las estadísticas censales actuales, pasadas o futuras para áreas pequeñas.
- Cuando se trata de la construcción de series temporales de estadísticas de áreas pequeñas es más útil extender la base geográfica actual hacia el pasado que tratar de preservar la base geográfica pasada. La encuesta que contempla el futuro de las tablas de consulta permitirá la construcción de tales series temporales.

En el IBGE, una vez concluido el sistema de agregación de sectores, para formar áreas de ponderación de la muestra del censo 2000, se deberá evaluar la posibilidad de compatibilizarlo (por agregación) con la base postal que está actualmente en revisión por el servicio de correos. La agregación de sectores reduce el problema de la confidencialidad pero no lo elimina.

3.3. Conducta ético-moral y procedimientos

Se encuentra en elaboración un **manual de conducta ético-moral**, que ya cuenta con una versión preliminar, dirigido a los funcionarios del IBGE, basado en varios textos legales destinados a los funcionarios públicos en general, y textos formales destinados a la generación de informaciones estadísticas y geocientíficas oficiales, entre los que destaca el titulado **Principios fundamentales de las estadísticas oficiales** (28° período de sesiones de la Comisión de Estadística de las Naciones Unidas, abril de 1994).

Al contrario de las orientaciones técnico-operacionales, que son de orden objetivo, los preceptos de conducta ético-moral son de orden subjetivo; por lo tanto, más efectivas serán esas

orientaciones que estos preceptos, sin perjuicio de que declararlos con toda claridad sea de la mayor importancia.

Cabe recordar que la **ética** se refiere al hecho de ajustar una vida común a principios filosóficos, a reflexionar sobre las razones de desear la justicia y la armonía, y los medios para alcanzarlas, pero sobre todo se trata de una teoría de los valores; y la **moral** se refiere a las costumbres, valores y normas de conducta concretos de una sociedad determinada, se refiere a la construcción de un conjunto de prescripciones destinadas a asegurar una vida en común, justa y armoniosa. Es decir, si por un lado la **ética** considera la acción humana desde un punto de vista valorativo y normativo, haciéndolo en un sentido genérico y abstracto, por otro, la **moral** establece los principios concretos que deben seguirse con miras a una vida en común.

En líneas generales, en los procesos de generación de la información figuran las conductas que deben observarse en las etapas de planificación, diseño, tratamiento, análisis e interpretación, mantención de los acervos, difusión, entre otros aspectos.

Dicho esto, hay que tener claro que un **manual de conducta ético-moral** actúa en tres vertientes complementarias, a saber: primero, postulando y orientando el cumplimiento riguroso e irrestricto de la ley estadística; segundo, postulando y orientando el cumplimiento estricto y riguroso de las orientaciones técnico-operacionales; tercero, indicando los caminos para tratar los casos de negligencia y faltas en las vertientes anteriores. Al enfocar estas tres vertientes, el manual privilegia un tratamiento de índole persuasivo, de carácter educativo (educación permanente), sin por ello dejar de enunciar los castigos que corresponden en situaciones de contumacia.

Además, le corresponde encauzar la institucionalidad de todo el proceso, estableciendo los principios para la constitución de una **comisión de ética**, como garante de la buena fe democrática de los procesos disciplinarios.

En cuanto al **manual técnico-operacional**, se analizaron aspectos relacionados con la utilización de determinados recursos computacionales, principalmente los relativos al uso del correo electrónico y la Internet. En muchos casos, es difícil separar el elemento operacional de los elementos legal y moral. Los dos tópicos abordados (correo electrónico, Internet) son ejemplos bien claros de dicha dificultad en la medida en que se mezclan reglas objetivas de cómo utilizar con parsimonia esos recursos y las facetas legales (limitaciones en cuanto al seguimiento de mensajes de correo y archivos anexados) y morales (derecho a la privacidad y seguimiento de los accesos a los datos) que presentan.

3.4. Acceso al acervo institucional de datos

En este acápite se presenta una propuesta analizada por el Grupo del Secreto sobre el acceso al acervo institucional de datos en cuanto a su relación con el **entorno de datos** del modelo de operación (letra D).

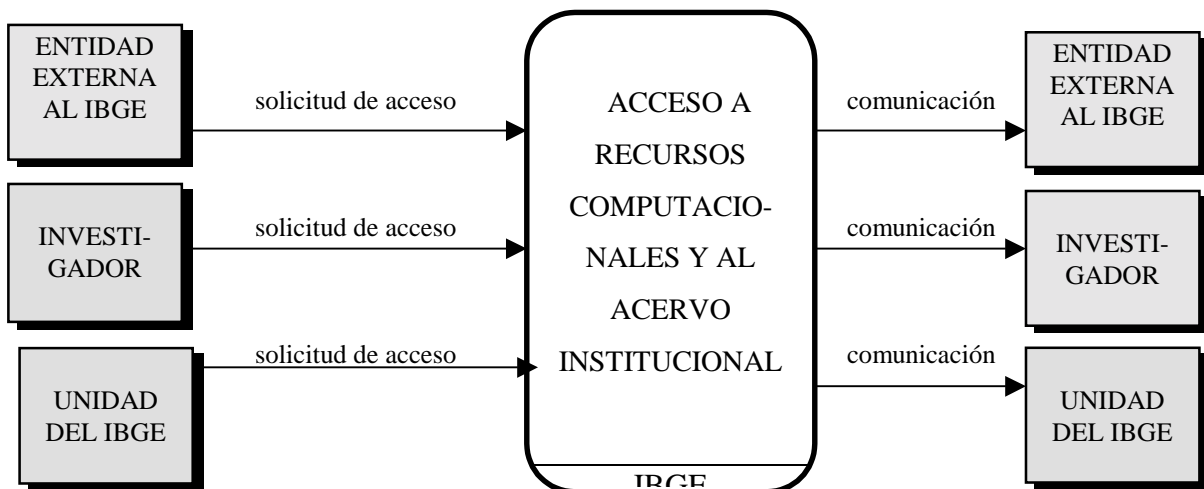
Las personas, sean ellas funcionarios del IBGE o de entidades externas al IBGE, pueden acceder al acervo institucional de microdatos del IBGE mediante una autorización que será otorgada por los encargados de las diferentes unidades del IBGE, sobre la base de reglas definidas por la institución.

Para acceder a los datos del acervo, no preparados para ser difundidos, hay que acceder previamente a los recursos computacionales del IBGE. Por ello, quien autoriza el acceso al acervo autoriza también el acceso a los recursos computacionales necesarios para posibilitar el acceso a los datos del acervo.

La apertura/restricción del acceso a los recursos y datos lo materializa la Dirección de Informática del IBGE mediante tal autorización.

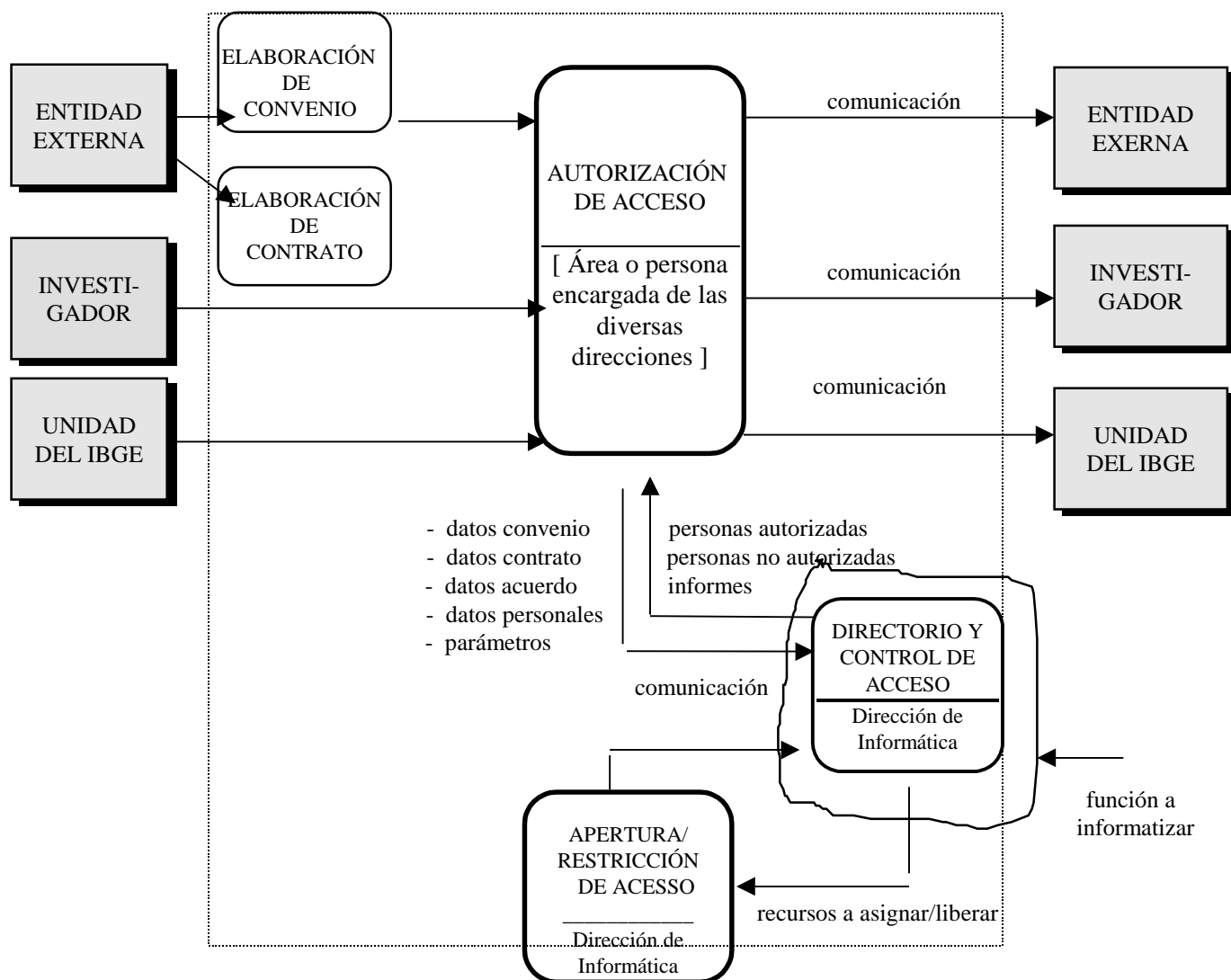
A continuación, se presentan dos diagramas con el fin de facilitar la visualización de las principales entidades y flujos de información del proceso de autorización, directorio y control de acceso con dos niveles de detalle diferentes.

PROCESO DE AUTORIZACIÓN, DIRECTORIO Y CONTROL DE ACCESO DIAGRAMA DE CONTEXTO



- personas autorizadas
 - personas no autorizadas
 - informe de seguimiento
-
- datos convenio
 - datos contratos
 - datos acuerdo
 - datos personales
 - parámetros

ACCESO A LOS RECURSOS COMPUTACIONALES Y AL ACERVO DE MICRODATOS DEL IBGE
PROCESO DE AUTORIZACIÓN, DIRECTORIO Y CONTROL DE ACCESO
DIAGRAMA NIVEL 0 (CERO)



En Bianchini y otros (1999) consta una descripción detallada de la propuesta de las condiciones para autorizar o cancelar el acceso a los datos, que involucra a todos los funcionarios activos, inactivos, personas vinculadas a entidades externas, personas físicas, vinculadas o no a personas jurídicas, por convenio o acuerdo. Además, constan las competencias en cuanto al acceso a recursos computacionales y datos del acervo institucional, con especificación de los datos necesarios para incluirlos o quitarlos del directorio.

4. Desafíos y perspectivas para el futuro

En vista de lo expuesto, se constata que hay múltiples avances y realizaciones sobre el tema de la confidencialidad de la información, pero que queda mucho por hacer, dada la complejidad del tema. Entre los principales productos se destacan:

- Elaborar la minuta de un proyecto de **ley estadística**, que se elevará a las autoridades competentes, que aglutine y actualice la legislación vigente, habida cuenta de la democracia y tecnología que imperan actualmente.
- Elaborar un **manual técnico-operacional** que contenga las reglas y procedimientos que deben seguir los diversos actores en cada uno de los puntos del modelo de operación.
- Concluir el **manual de conducta ético-moral** que contenga los principios que deben seguir los funcionarios del IBGE en su calidad de actores situados en los diversos puntos del modelo de operación.

En la elaboración de esos productos se trabaja en la implementación descentralizada en que se asignan recursos a proyectos con miras a las siguientes actividades:

- Investigación, absorción, verificación, evaluación y difusión de **métodos y sistemas automatizados de protección**. Este proyecto debe contemplar también la **creación de cursos en el programa anual de capacitación** sobre el secreto y sus efectos en los métodos y procesos de encuesta.
- Implantación de procedimientos que garanticen la completa **seguridad de acceso a los datos** del acervo institucional.
- **Creación de espacios adecuados para accesos especiales** al acervo por personas autorizadas.
- Adquisición de los equipos necesarios para la **adecuada destrucción de material** con informaciones individuales.

Cabe formular tres observaciones sobre la mantención del secreto garantizado a los informantes, un aspecto esencial en el presente y futuro de los organismos de estadística, de modo que haya una mayor y mejor expedición en lo que se estará haciendo en los próximos años:

- Las medidas adoptadas o por adoptar exigen gran cantidad de recursos humanos, pero no sólo eso, pues es un conjunto de actividades que implica **costos** que hay que asumir seriamente.
- Las acciones presentes y futuras implican cambios en la **cultura** institucional, lo que exige actuar en todos los niveles funcionales, sin excepción, involucrados en las diversas etapas de generación de la información;
- Es preciso incorporar en la **planificación de cada encuesta** los procedimientos de protección de la información, desde la recopilación hasta la difusión. El compromiso de todos los departamentos y áreas de la organización relacionados con el tema, mediante la capacitación y el planteamiento metódico de la situación vigente en cada una de las encuestas.

En vista de lo expuesto, se debe entender que los resultados, dados los atrasos temporales y espaciales que lamentablemente existen en la institución, sólo se harán sentir a mediano y largo plazo por más urgentes y esenciales que sean.

Por último, considerando que el tema merece una dedicación constante, se creó recientemente el Comité de la Confidencialidad, que pasa a ser el lugar de reflexión para las cuestiones relacionadas con el secreto, con las atribuciones siguientes:

- a) **proponer soluciones** para cuestiones relacionadas con el secreto de la información de naturaleza estadística y geocientífica, individual o identificada, recopilada, producida, almacenada y difundida por el IBGE;
- b) **sopesar** las solicitudes de acceso a la información confidencial, dando su parecer sobre la autorización y condiciones de uso; y
- c) **evaluar** sistemáticamente el cumplimiento de la promesa de asegurar la mantención del secreto de la información confidencial.

Bibliografia

- Anzanello, E. (1999), *Modelo de operação para pesquisas estatísticas e sigilo de informações*, Rio de Janeiro, Dirección de Informática, Instituto Brasileño de Geografía y Estadística (IBGE).
- Bethlehem, J.G., W.J. Kellere y J. Pannekoek (1990), "Disclosure control of microdata", *Journal of the American Statistical Association*, N° 85.
- Bianchini, Z. M. y otros (1999), *Grupo de Sigilo: atividades realizadas, desafios e perspectivas para o futuro*, Rio de Janeiro, Instituto Brasileño de Geografía y Estadística (IBGE).
- Bianchini, Z.M. (1995), *Considerações sobre o sigilo das informações. Mercosur: sinopsis estadística*, vol. 2, Buenos Aires.
- Burshtein, G. y G. Lang (1999), "Administration and policy of statistical data confidentiality in Israel", *Joint ECE/Eurostat Work Session on Statistical Data Confidentiality*, Working Paper N° 22, Atenas.
- Carvalho, J. (1999), *Sigilo: Legislação básica*, Rio de Janeiro, Instituto Brasileño de Geografía y Estadística (IBGE)/ Procuraduría General.
- Cecil, J.S. (1993), "Confidentiality legislation and the United States Federal Statistical System", *Journal of Official Statistics*, vol. 9, N° 2.
- Comisión de Estadística de las Naciones Unidas (1994), *Principios fundamentales de las estadísticas oficiales*, Nueva York.
- Cook, L. (1998), *Protocols for Official Statistics*, Nueva Zelandia.
- Cox, L.H. (1994), "Protecting confidentiality in establishment surveys", *Survey Methods for Businesses, Farms and Institutions*, Brenda Cox (comp.), Nueva York, John Wiley and Sons.
- Defays, D. y M.N. Anwar, (1998), "Masking microdata using micro-aggregation", *Journal of Official Statistics*, vol. 14, N° 4.
- Duncan, G.T. (2000), "Optimal disclosure limitation strategy in statistical databases: deterring tracker attacks through additive noise", *Journal of the American Statistical Association*, vol. 95, N° 451.
- Elliot, M. y A. Dale (1999), "Scenarios of attack: the data intruder's perspective on statistical disclosure risk", *Netherlands Official Statistics*, vol. 14, número especial.
- Fienberg, S.E. y L.C.R.J. Willenborg (1998), "Introduction to the special issue: disclosure limitation methods for protecting the confidentiality of statistical data", *Journal of Official Statistics*, vol. 14, N° 4.
- Fienberg, S.E. y U.E. Makov (1998), "Confidentiality, uniqueness, and disclosure limitation for categorical data", *Journal of Official Statistics*, vol. 14, N° 4.
- Fuller, W.A. (1993), "Masking procedures for disclosure limitation", *Journal of Official Statistics*, vol. 9, N° 2.
- Gouweleeuw, J.M. y otros (1998), "Post randomisation for statistical disclosure control: theory and implementation", *Journal of Official Statistics*, vol. 14, N° 4.
- Greenberg, B. (1990), "Disclosure avoidance research at the Bureau of the Census", *Proceedings of Sixth Annual Research Conference*, Washington, D.C., Oficina del Censo.
- Hurkens, C.A.J. y S.R. Tiourine (1998), "Models and methods for the microdata protection problem", *Journal of Official Statistics*, vol. 14, N° 4.

- IBGE (Instituto Brasileiro de Geografia y Estadística) (1996), *Requisitos técnicos para uma lei sobre a produção e o uso de informações estatísticas*, Rio de Janeiro.
- INDEC (Instituto Nacional de Estadística y Censos) (1988), *Manual sobre secreto estadístico*, vol. 2, Buenos Aires.
- Jabine, T.B. (1993), “Statistical disclosure limitation practices of United States Statistical Agencies”, *Journal of Official Statistics*, vol. 9, N° 2.
- Keller, W.J. y J.G. Bethlehem (1992), “Disclosure protection of microdata: problems and solutions”, *Statistica Neerlandica*, N° 46.
- Keller-McNulty, S. y E.A. Unger (1998), “Database system prototype for remote access to information based on confidential data”, *Journal of Official Statistics*, vol. 14, N° 4.
- Kirkendall, N.J. y otros (1994), “Report on Statistical Limitation Methodology (Subcommittee on Disclosure Limitation Methodology)”, Statistical Policy Working Paper, N° 22, Subcomité de Metodología sobre la Limitación de la Revelación.
- Kooiman, P., J. Nobele y L. Willenborg (1999), “Statistical data protection at Statistics Netherlands”, *Netherlands Official Statistics*, vol. 14, número especial.
- Lang, G. (1999), “Statistical confidentiality and the French committee of statistical confidentiality concerning enterprises”, *Joint ECE/Eurostat Work Session on Statistical Data Confidentiality*, Working Paper, N° 20, Atenas.
- Marsh, C., A. Dale y C. J. Skinner (1994), “Safe data versus safe settings: access to microdata from the British Census”, *International Statistical Review*, vol. 62, N° 1.
- McLenaghan, J.B. (1999), “The UN’s fundamental principles of official statistics issues of coverage and application”, *Working Session On Best Practices in Statistics*, Singapur.
- Michael, J.A. y otros (1978), “Report on Statistical Disclosure-Avoidance Techniques (Subcommittee on Disclosure-Avoidance Techniques Federal Committee on Statistical Methodology)”, Statistical Policy Working Paper, N° 2, Subcomité de Metodología sobre la Limitación de la Revelación.
- Nordholt, E.S. (1999), “Statistical disclosure control of Statistical Netherlands employment and earnings data”, *Netherlands Official Statistics*, vol. 14, número especial.
- Openshaw, S., O. Duke-Williams e P. Rees (1997), “Measuring Confidentiality in Census Data”, Working Paper, N° 97/8.
- Pannekoek, J. y T. Waal (1998), “Synthetic and combined estimators in statistical disclosure control”, *Journal of Official Statistics*, vol. 14, N° 4.
- Rees, P. y O. Duke-Williams (1998), *Data Privacy Considerations: The Differencing Problem in Statistical Disclosure*, Londres.
- _____ (1997), *Experiments with and Recommendations for the Creation and Release of Small Statistics from National Censuses*, Londres, Universidad de Leeds.
- Robertson, D.A. (1993), “Cell suppression at Statistics Canada”, *Proceedings of the Bureau of the 1993 Census Annual Research Conference*, Washington, D.C., Oficina del Censo.
- Samuels, S.M. (1998), “A Bayesian, species-sampling-inspired approach to the unique problem in microdata disclosure risk assessment”, *Journal of Official Statistics*, vol. 14, N° 4.
- Senra, N.C. (1999a), *Garantía de Sigilo, a deontología do estatístico: sua conduta ético-moral*, Rio de Janeiro, Instituto Brasileiro de Geografia y Estadística (IBGE).
- _____ (1999b), *Reflexões sobre a questão do sigilo*, Rio de Janeiro, Instituto Brasileiro de Geografia y Estadística (IBGE).

- Servicio Estadístico del Gobierno del Reino Unido (1995), *Official Statistics: Code of Practice*, Londres.
- Silva, P.L.N. (1999), *Sigilo Estatístico e disseminação de informações*, Rio de Janeiro, Departamento de Metodología, Instituto Brasileiro de Geografia y Estadística (IBGE).
- _____. (1988), *O sigilo das informações estatísticas: idéias para reflexão*, Rio de Janeiro, Instituto Brasileiro de Geografia y Estadística (IBGE).
- Skinner, C.J. y D.J. Holmes (1998), "Estimating the re-identification risk for record in microdata", *Journal of Official Statistics*, vol. 14, N° 4.
- Skinner, C.J. y otros (1990), "Disclosure avoidance for census microdata in Great Britain", *Proceedings of the Sixth Annual Research Conference*, Washington, D.C., Bureau of the Census.
- Spieker, F. (1999), "Access to confidential data in an integrated statistical system", *Joint ECE/Eurostat Work Session on Statistical Data Confidentiality*, Working Paper, N° 18, Atenas.
- Thorogood, D. (1999a), "Statistical confidentiality and the European level", *Joint ECE/Eurostat Work Session on Statistical Data Confidentiality*, Working Paper, N° 16, Atenas.
- _____. (1999b), "Protecting the confidentiality of Eurostat statistical outputs", *Netherlands Official Statistics*, vol. 14, número especial.
- Unión Europea (1999), *El secreto estadístico en las legislaciones nacionales de los Estados miembros de la Unión Europea*, Madrid.
- Vasconcellos, M.T.L. (1996), *A disseminação dos dados e a proteção do informante: idéias básicas para a definição de uma política institucional*, Rio de Janeiro, Departamento de Metodología, Instituto Brasileiro de Geografia y Estadística (IBGE).
- Waal, T. y L. Willenborg (1999a), "Information loss through global recoding and local suppression", *Netherlands Official Statistics*, vol. 14, número especial.
- _____. (1999b), "Exact disclosure in a super-table", *Netherlands Official Statistics*, vol. 14, número especial.
- _____. (1998), "Optimal local suppression in microdata", *Journal of Official Statistics*, vol. 14, N° 4.
- Willenborg, L.C.R.J. (1993), "Discussion: statistical disclosure limitation", *Journal of Official Statistics*, vol. 9, N° 2.
- Willenborg, L.C.R.J. y T. Waal (1996), "Statistical disclosure control in practice", *Lecture Notes in Statistics*, N° 111, Nueva York, Springer-Verlag.
- Willenborg, L.C.R.J., R.J. Mokken y J. Pannekoek (1990), "Microdata and disclosure risks", *Proceedings of the Sixth Annual Research Conference*, Washington, D.C., Oficina del Censo.
- Zaslavsky, A.M. y N.J. Horton (1998), "Balancing disclosure risk against the loss of nonpublication", *Journal of Official Statistics*, vol. 14, N° 4.
- Zayatz, L., P. Massell y P. Steel (1999), "Disclosure limitation practices and research at the U.S. Census Bureau", *Netherlands Official Statistics*, vol. 14, número especial.